PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 : C12Q 1/68	A1	(11) International Publication Number: WO 93/23564
		(43) International Publication Date: 25 November 1993 (25.11.93)

(21) International Application Number: PCT/EP93/01205

(22) International Filing Date: 12 May 1993 (12.05.93)

(30) Priority data:
9210168.2 12 May 1992 (12.05.92) GB

(71) Applicant *(for all designated States except US)*: CEMUBIO-TEKNIK AB [SE/SE]; Banergatan 21, S-752 37 Uppsala (SE).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : UHLEN, Mathias [SE/SE]; Kvarnbogatan 30, S-752 39 Uppsala (SE). NYREN, Pål [SE/SE]; Riksrådsvägen 67, S-128 39 Skarpnok (SE).

(74) Agents: HOLMES, Michael, John et al.; Frank B. Dehn & Co., Imperial House, 15-19 Kingsway, London WC2B 6UZ (GB).

(81) Designated States: AU, CA, JP, NO, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.

(54) Title: METHOD OF SEQUENCING DNA

(57) Abstract

The method provides a method of identifying a base at a target position in a single-stranded DNA sequence wherein an extension primer, which hybridises to the sample DNA immediately adjacent to the target position is provided and the sample DNA and extension primer are subjected to a polymerase reaction in the presence of a deoxynucleotide or dideoxynucleotide whereby the deoxynucleotide or dideoxynucleotide will only become incorporated and release pyrophosphate if it is complementary to the base in the target position, any release of pyrophosphate being identified enzymically, different deoxynucleotides or dideoxynucleotides being added either to separate aliguots of sample-primer mixture or successively to the same sample-primer mixture and subjected to the polymerase reaction to indicate which deoxynucleotide or dideoxynucleotide is incorporated.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	**FR**	France	**MR**	Mauritania
AU	Australia	**GA**	Gabon	**MW**	Malawi
BB	Barbados	**GB**	United Kingdom	**NL**	Netherlands
BE	Belgium	**GN**	Guinea	**NO**	Norway
BF	Burkina Faso	**GR**	Greece	**NZ**	New Zealand
BG	Bulgaria	**HU**	Hungary	**PL**	Poland
BJ	Benin	**IE**	Ireland	**PT**	Portugal
BR	Brazil	**IT**	Italy	**RO**	Romania
CA	Canada	**JP**	Japan	**RU**	Russian Federation
CF	Central African Republic	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CG	Congo			**SE**	Sweden
CH	Switzerland	**KR**	Republic of Korea	**SK**	Slovak Republic
CI	Côte d'Ivoire	**KZ**	Kazakhstan	**SN**	Senegal
CM	Cameroon	**LI**	Liechtenstein	**SU**	Soviet Union
CS	Czechoslovakia	**LK**	Sri Lanka	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**MC**	Monaco	**UA**	Ukraine
DK	Denmark	**MG**	Madagascar	**US**	United States of America
ES	Spain	**ML**	Mali	**VN**	Viet Nam
FI	Finland	**MN**	Mongolia		

Method of sequencing DNA

This invention relates to a method of sequencing DNA.

DNA sequencing is an essential tool in molecular genetic analysis. The ability to determine DNA nucleotide sequences has become increasingly important as efforts have commenced to determine the sequences of the large genomes of humans and other higher organisms. The two most commonly used methods for DNA sequencing are the enzymatic chain-termination method of Sanger et al. (1) and the chemical cleavage technique by Maxam and Gilbert (2). Both methods rely on gel electrophoresis to resolve, according to their size, DNA fragments produced from a larger DNA segment. Since the electrophoresis step as well as the subsequent detection of the separated DNA-fragments are cumbersome procedures, a great effort has been made to automate these steps (3,4). However, despite the fact that automated electrophoresis units are commercially available, electrophoresis is not well suited for large-scale genome projects or clinical sequencing where relatively cost-effective units with high throughput are needed. Thus, the need for non-electrophoretic methods for sequencing is great and recently several alternative strategies have been described, such as scanning tunnel electron microscopy (5), sequencing by hybridization (6, 7) and single molecule detection (8), to overcome the disadvantages of electrophoresis.

Techniques enabling the rapid detection of a single DNA base change are also important tools for genetic analysis. In many cases detection of a single base or a few bases would be a great help in genetic analysis since several genetic diseases and certain cancers are related to minor mutations (9, 10). Recently, a mini-

sequencing protocol based on a solid phase principle was described (11, 12). The incorporation of a radiolabeled nucleotide was measured and used for analysis of the three-allelic polymorphism of the human apolipoprotein E gene. However, radioactive methods are not well suited for routine clinical applications and the development of a simple non-radioactive method for rapid DNA sequence analysis is of considerably interest.

There is a need for a simple and rapid method for identifying a base in a target position and for DNA sequencing which avoids the need for electrophoresis and the use of harmful radiolabels.

The present invention is based on the concept of detecting inorganic pyrophosphate (ppi) which is released during a polymerase reaction. As each nucleotide is added to a growing nucleic acid strand during a polymerase reaction, a pyrophosphate molecule is released. We have found that pyrophosphate released under these conditions can be detected enzymically e.g. by the generation of light in the luciferase-luciferin reaction.

The present invention provides a method of identifying a base at a target position in a single-stranded DNA sequence wherein an extension primer, which hybridises to the sample DNA immediately adjacent to the target position is provided and the sample DNA and extension primer are subjected to a polymerase reaction in the presence of a deoxynucleotide or dideoxynucleotide whereby the deoxynucleotide or dideoxynucleotide will only become incorporated and release pyrophosphate if it is complementary to the base in the target position, any release of pyrophosphate being identified enzymically, different deoxynucleotides or dideoxynucleotides being added either to separate aliquots of sample-primer mixture or successively to the same sample-primer mixture and subjected to the polymerase reaction to indicate which deoxynucleotide or

dideoxynucleotide is incorporated.

The term dideoxynucleotide as used herein includes all 2'-deoxynucleotides in which the 3'-hydroxyl group is absent or modified and thus, while able to be added to the primer in the presence of the polymerase, is unable to enter into a subsequent polymerisation reaction.

It is preferred to use luciferase and luciferin in combination to identify the release of pyrophosphate since the amount of light generated is substantially proportional to the amount of pyrophosphate released which, in turn, is directly proportional to the amount of base incorporated. The amount of light can readily be estimated by a suitable light sensitive device such as a luminometer.

Where ATP is present in the reaction mixture after chain extension, for example as an impurity or as a contaminant of dATP added as the source of the base to be incorporated, it will interfere in the pyrophosphate luciferin system and given an incorrect luminescence reading. We have found it advantageous, therefore, to remove ATP from such solutions prior to addition of luciferase/luciferin. This can be achieved by contacting the solution with an immobilised enzyme which converts ATP into a product which is no longer a substrate for luciferase. Such enzymes include, in particular, apyrase which converts the ATP to AMP and two molecules of phosphate. The immobilised enzyme may then be removed prior to the luminescent reaction. It is particularly convenient to use magnetic beads such as Dynabeads® (sold by Dynal AS, Oslo, Norway) as the solid support due to the ease with which such beads can be removed from contact with the solution using a magnet.

In order to repeat the method cyclically and thereby sequence the sample DNA and, also to aid separation of the single stranded sample DNA from its complementary strand, it is desirable that the sample

DNA is immobilised or provided with means for attachment to a solid support. Moreover, the amount of sample DNA available may be small and it will therefore be desirable to amplify the sample DNA before carrying out the method according to the invention.

Preferably, therefore, the sample DNA is amplified, for example in vitro by PCR or Self Sustained Sequence Replication (3SR) or in vivo using a vector and, if desired, in vitro and in vivo amplification may be used in combination. Whichever method of amplification is used it is desirable that the amplified DNA becomes immobilised or is provided with means for attachment to a solid support. For example, a PCR primer may be immobilised or be provided with means for attachment to a solid support. Also, a vector may comprise means for attachment to a solid support adjacent the site of insertion of the sample DNA such that the amplified sample DNA and the means for attachment may be excised together.

Immobilisation of the amplified DNA may take place as part of PCR amplification itself, as where one or more primers are attached to a support, or alternatively one or more of the PCR primers may carry a functional group permitting subsequent immobilisation, eg. a biotin or thiol group. Immobilisation by the 5' end of a primer allows the strand of DNA emanating from that primer to be attached to a solid support and have its 3' end remote from the support and available for subsequent hybridisation with the extension primer and chain extension by polymerase.

The solid support may conveniently take the form of microtitre wells, which are advantageously in the conventional 8 x 12 format, or dipsticks which may be made of polystyrene activated to bind the primer DNA (K Almer, Doctoral Theses, Royal Institute of Technology, Stockholm, Sweden, 1988). The support may also comprise particles, fibres or capillaries made, for example, of

agarose, cellulose, alginate, Teflon or polystyrene. Magnetic particles eg the superparamagnetic beads produced by Dynal AS (Oslo, Norway) are a preferred support since they can be readily isolated from a reaction mixture yet have superior reaction kinetics over many other forms of support.

The solid support may carry functional groups such as hydroxyl, carboxyl, aldehyde or amino groups, or other moieties such as avidin or streptavidin, for the attachment of primers. These may in general be provided by treating the support to provide a surface coating of a polymer carrying one of such functional groups, e.g. polyurethane together with a polyglycol to provide hydroxyl groups, or a cellulose derivative to provide hydroxyl groups, a polymer or copolymer of acrylic acid or methacrylic acid to provide carboxyl groups or an aminoalkylated polymer to provide amino groups. US Patent No. 4654267 describes the introduction of many such surface coatings.

The assay technique is very simple and rapid, thus making it easy to automate by using a robot apparatus where a large number of samples may be rapidly analysed. Since the preferred detection and quantification is based on a luminometric reaction which can be easily followed spectrophotometrically.

The target DNA may be cDNA synthesised from RNA in the sample and the method of the invention is thus applicable to diagnosis on the basis of characteristic RNA. Such preliminary synthesis can be carried out by a preliminary treatment with a reverse transcriptase, conveniently in the same system of buffers and bases of subsequent PCR steps if used. Since the PCR procedure requires heating to effect strand separation, the reverse transcriptase will be inactivated in the first PCR cycle. When mRNA is the sample nucleic acid, it may be advantageous to submit the initial sample, e.g. a serum sample, to treatment with an immobilised polydT

oligonucleotide in order to retrieve all mRNA via the terminal polyA sequences thereof. Alternatively, a specific oligonucleotide sequence may be used to retrieve the RNA via a specific RNA sequence. The oligonucleotide can then serve as a primer for cDNA synthesis, as described in International Patent Application PCT/89EP/00304.

Advantageously, the extension primer is sufficiently large to provide appropriate hybridisation with the sequence immediately 5' of the target position, yet still reasonably short in order to avoid unnecessary chemical synthesis. It will be clear to persons skilled in the art that the size of the extension primer and the stability of hybridisation will be dependent to some degree on the ratio of A-T to C-G base pairings, since more hydrogen bonding is available in a C-G pairing. Also, the skilled person will consider the degree of homology between the extension primer to other parts of the amplified sequence and choose the degree of stringency accordingly. Guidance for such routine experimentation can be found in the literature, for example, Molecular Cloning: a laboratory manual by Sambrook, J., Fritsch E.F. and Maniatis, T. (1989). The extension primer is preferably added before the sample is divided into four aliquots although it may be added separately to each aliquot. It should be noted that the extension primer may be identical with the PCR primer but preferably it is different, to introduce a further element of specificity into the system.

The polymerase reaction in each aliquot in the presence of the extension primer and a deoxynucleotide is carried out using a polymerase which will incorporate dideoxynucleotides, e.g. T7 polymerase, Klenow or Sequenase Ver. 2.0 (USB U.S.A.). However, it is known that many polymerases have a proof-reading or error checking ability and that 3' ends available for chain extension are sometimes digested by one or more

nucleotides. If such digestion occurs in the method according to the invention the level of background noise increases. In order to avoid this problem it is preferable to use a nonproof-reading polymerase, eg T7 polymerase or Sequenase. Otherwise it is desirable to add to each aliquot fluoride ions or nucleotide monophosphates which suppress 3' digestion by polymerase.

In many diagnostic applications, for example genetic testing for carriers of inherited disease, the sample will contain heterozygous material, that is half the DNA will have one nucleotide at the target position and the other half will have another nucleotide. Thus if four aliquots are used in a preferred method according to the invention, two will show a positive signal and two will show half the positive signal. It will be seen therefore that it is desirable to quantitatively determine the amount of label detected in each sample. Also, it will be appreciated that if two or more of the same base are adjacent the 3'-end of the primer a larger signal will be produced. In the case of a homozygous sample it will be clear that there will be three negative and one positive signal when the same is in four aliquots.

It will be appreciated that when the target base immediately 3'- of the primer has an identical base 3'-thereto, and the polymerisation is effected with a deoxynucleotide (rather than a dideoxynucleotide) the extension reaction will add two bases at the same time and indeed any sequence of successive identical bases in the sample will lead to simultaneous incorporation of corresponding bases into the primer. However, the amount of pyrophosphate liberated will clearly be proportional to the number of incorporated bases so that there is no difficulty in detecting such repetitions.

Since the primer is extended by a single base by the procedure described above (or a sequence of

identical bases), the extended primer can serve in exactly the same way in a repeated procedure to determine the next base in the sequence, thus permitting the whole sample to be sequenced. Immobilisation of the sample and hybridised primer permits washing to separate unwanted deoxynucleotides before proceeding to the next step.

The present invention provides two principal methods of sequencing immobilised DNA.

A. The invention provides a first method of sequencing sample DNA wherein the sample DNA is subjected to amplification; the amplified DNA is immobilised and then subjected to strand separation, the non-immobilised strand being removed and an extension primer is provided, which primer hybridises to the immobilised DNA immediately adjacent that portion of the DNA to be sequenced; each of four aliquots of the immobilised single stranded DNA is then subjected to a polymerase reaction in the presence of a deoxynucleotide, each aliquot using a different deoxynucleotide whereby only the deoxynucleotide complementary to the base in the target position becomes incorporated; pyrophosphate released by base incorporation being identified; the immobilised sample and primer then being separated from the reaction solution and the incorporated base added to the unreacted aliquots of sample/primer under polymerising conditions to extend the primer in all the aliquots by the said incorporated base and the immobilised sample/primer then being separated from the reaction solution, the process being repeated to sequence the sample DNA.

B. The invention also provides a second method of sequencing sample DNA wherein the sample DNA is subjected to amplification; the amplified DNA is immobilised and then subjected to strand separation, the non-immobilised strand being removed and an extension primer is provided, which primer hybridises to the

immobilised DNA immediately adjacent that portion of the DNA to be sequenced; immobilised single stranded DNA is then subjected to a polymerase reaction in the presence of a first deoxynucleotide, and the extent of pyrophosphate release is determined, where necessary the immobilised sample and primer being separated from the reaction mixture and the reaction being repeated by successive addition of a second, third and fourth deoxynucleotide until a positive release of pyrophosphate indicates incorporation of a particular deoxynucleotide into the primer, whereupon the procedure is repeated to extend the primer one base at a time and to determine the base which is immediately 3'- of the extended primer at each stage.

It is desirable that when PCR is used in initial amplification its effectiveness is assessed, e.g. to determine whether or not sufficient DNA has been formed to give clear results with a relatively low level of background. Various tests are known in the art but we prefer to use the solid phase approach we described earlier for detection of immobilized amplified nucleic acids, designated DIANA (PCT/EP90/00454), which has been used for example in its preferred embodiment in the colorimetric detection of in vitro amplified DNA. The assay is based on the use of a biotinylated or otherwise functionalised PCR primer, which is used to capture in vitro amplified material on, for example, streptavidin-coated magnetic beads. The other PCR primer contains a "handle", such as a lac operator sequence, allowing colorimetric detection of the captured DNA using a LacI repressor-β-galactosidase fusion protein. (Wahlberg, J., Lundeberg, J., Hultman, T. and Uhlén, M. (1990) "General colorimetric method for DNA diagnostics allowing direct solid-phase genomic sequencing of the positive samples". Proc. Natl. Acad. U.S.A. 87 6569-6573). The preferred form of the qualitative DIANA assay combines the advantages of the PCR method with the high specificity

and stability of the biotin-streptavidin system and the simplicity of a colorimetric detection based on β-galactosidase. The strong interaction between biotin and streptavidin ($K_d=10^{-15}$ M^{-1}) accentuates the efficiency of the system. The magnetic beads as solid support ensure that no centrifugations, filtrations or precipitations are needed (T. Hultman, S. Ståhl, E. Hornes and M. Uhlén Nucl. Acids Res. 17, 4937 (1989)). However, it is preferred in the method according to the present invention to use the same PCR primer both as the means of immobilisation and for the incorporation of the lac operator sequence.

A number of proteins are known which bind to specific DNA sequences and are often involved in genetic processes such as switching operons on and off. One such protein is the lac repressor LacI which reacts with the lac operator (lacOP) to inhibit transcription. Thus, if the recognition site is the DNA sequence lacOP, the label can be attached via the protein LacI. It is particularly convenient to devise a fusion protein of a DNA binding protein such as LacI with a further protein which can be subsequently used for detection for example using methods based on colour fluorescence or chemiluminescence. Examples of such proteins are β-galactosidase, alkaline phosphatase and peroxidase.

It is preferred to use as a label a LacI repressor-β-galactosidase fusion protein which recognises a 21 base pair lac operator sequence introduced at the end of the amplified DNA. The lac operator sequence may be introduced for example by one of the PCR primers if used, preferably the immobilised primer, or the sequence may be in an amplification vector in a suitable position for excision with the amplified sample DNA. The fusion protein will bind to the LacOP sequence of the DNA and the addition of ONPG (ortho-nitrophenyl-β-D-galactoside) will lead to a colour formation which can be assessed spectrophotometrically. Use of this fusion protein and

ONPG allows for a fast simple colorimetric assay which does not have the safety problems associated with using radiolabels. IPTG (n-isopropyl-β-D-thiogalactopyranoside), for example, can be added to release the fusion protein from the DNA.

Two-stage PCR (using nested primers), as described in our co-pending application PCT/EP90/00454, may be used to enhance the signal to noise ratio and thereby increase the sensitivity of the method according to the invention. By such preliminary amplification, the concentration of target DNA is greatly increased with respect to other DNA which may be present in the sample and a second-stage amplification with at least one primer specific to a different sequence of the target DNA significantly enhances the signal due to the target DNA relative to the 'background noise'.

Regardless of whether one-stage or two stage PCR is performed, the efficiency of the PCR is not critical since the invention relies on the distinct difference different from the aliquots. However, as mentioned above, it is preferred to run an initial qualitative DIANA as a check for the presence or absence of amplified DNA.

Any suitable polymerase may be used, although it is preferred to use a thermophilic enzyme such as Taq polymerase to permit the repeated temperature cycling without having to add further polymerase, e.g. Klenow fragment, in each cycle of PCR.

PCR has been discussed above as a preferred method of initially amplifying target DNA although the skilled person will appreciate that other methods may be used instead of in combination with PCR. A recent development in amplification techniques which does not require temperature cycling or use of a thermostable polymerase is Self Sustained Sequence Replication (3SR). 3SR is modelled on retroviral replication and may be used for amplification (see for example Gingeras, T.R.

et al PNAS (USA) 87:1874-1878 and Gingeras, T.R. et al PCR Methods and Applications Vol. 1, pp 25-33).

As indicated above, the method can be applied to identifying the release of pyrophosphate when dideoxynucleotide residues are incorporated into the end of a DNA chain. Our co-pending United Kingdom application No. 9210176.5 of even date entitled "Chemical" Method" relates to a method of identification of the base in a single target position in a DNA sequence (mini-sequencing) wherein sample DNA is subjected to amplification; the amplified DNA is immobilised and then subjected to strand separation, the non-immobilised strand being removed and an extension primer, which hybridises to the immobilised DNA immediately adjacent to the target position, is provided; each of four aliquots of the immobilised single stranded DNA is then subjected to a polymerase reaction in the presence of a dideoxynucleotide, each aliquot using a different dideoxynucleotide whereby only the dideoxynucleotide complementary to the base in the target position becomes incorporated; the four aliquots are then subjected to extension in the presence of all four deoxynucleotides, whereby in each aliquot the DNA which has not reacted with the dideoxynucleotide is extended to form double stranded DNA while the dideoxy-blocked DNA remains as single stranded DNA; followed by identification of the double stranded and/or single stranded DNA to indicate which dideoxynucleotide was incorporated and hence which base was present in the target position. Clearly, the release of pyrophosphate in the chain terminating dideoxynucleotide reaction will indicate which base was incorporated but the relatively large amount of pyrophosphate released in the subsequent deoxynucleotide primer extension reactions (so-called chase reactions) gives a much larger signal and is thus more sensitive.

It will usually be desirable to run a control with

no dideoxynucleotides and a 'zero control' containing a mixture of all four dideoxynucleotides.

Our above copending Application No. 9210176.5 defines the term 'dideoxynucleotide' as including 3'-protected 2'-deoxynucleotides which act in the same way by preventing further chain extension. However, if the 3' protecting group is removable, for example by hydrolysis, then chain extension (by a single base) may be followed by unblocking at the 3' position, leaving the extended chain ready for a further extension reaction. In this way, chain extension can proceed one position at a time without the complication which arises with a sequence of identical bases, as discussed above. Thus, the methods A and B referred to above can be modified whereby the base added at each stage is a 3'-protected 2'-deoxynucleotide and after the base has been added (and the light emission detected), the 3'-blocking group is removed to permit a further 3'-protected - 2' deoxynucleotide to be added. Suitable protecting groups include acyl groups such as alkanol grouops e.g. acetyl or indeed any hydroxyl protecting groups known in the art, for example as described in Protective Groups in Organic Chemistry, JFW McOnie, Plenum Press, 1973.

The invention, in the above embodiment, provides a simple and rapid method for detection of single base changes. In a preferred format it successfully combines two techniques: solid-phase technology (DNA bound to magnetic beads) and an Enzymic Luminometric Detection Assay (ELIDA). The method can be used to both identify and quantitate selectively amplified DNA fragments. It can also be used for detection of single base substitutions and for estimation of the heterozygosity index for an amplified polymorphic gene fragment. This means that the method can be used to screen for rare point mutations responsible for both acquired and inherited diseases, identify DNA polymorphisms, and even differentiate between drug-resistant and drug-sensitive

strains of viruses or bacteria without the need for centrifugations, filtrations, extractions or electrophoresis. The simplicity of the method renders it suitable for many medical (routine analysis in a wide range of inherited disorders) and commercial applications.

The positive experimental results presented below clearly show the method is applicable to an on-line automatic non-electrophoretic solid phase DNA sequencing approach, with step-wise incorporation of single deoxynucleotides. After amplification, immobilization on magnetic beads, melting to yield single-stranded DNA and annealing of the primer, the template/primer-fragment is used in a repeated cycle of dNTP incubation and washing in four aliquots. Samples are continuously used in the ELIDA. As the synthesis of DNA is accompanied by release of inorganic pyrophosphate (ppi) in an amount equal to the amount of nucleotide incorporated, signals in the ELIDA are observed only when complementary bases are incorporated. In the single-base extension (mini-sequencing) reactions presented below, the possibility to incorporate a single complementary nucleotide with little or no background of noncomplementary nucleotide is shown. Incorporation of a single dNTP into DNA has also been reported earlier (17, 18). Increasing the amount of DNA would make it possible to detect the ppi formed during a single base incorporation. Due to the ability of the method to determine ppi quantitatively, it is possible to distinguish incorporation of a single base from two or several simultaneous incorporations. Since the DNA template is preferably obtained by PCR, it is relatively straight forward to increase the amount of DNA needed for such an assay.

As mentioned above our results open the possibility for a novel approach for large-scale non-electrophoretic solid phase DNA sequencing, which allows for continuous

determination of the progress of the polymerisation reaction with time. For the success of such an approach there is a need for high efficiency of the DNA polymerase due to the rapid increase of background signal if templates accumulate which are not "in phase". Our repeated mini-sequencing studies clearly demonstrate low background. The new approach has several advantages as compared to standard sequencing methods. Firstly, the method is suitable for handling of multiple samples in parallel. Secondly, relatively cost-effective instruments can be envisioned. In addition, the method avoids the use of electrophoresis and thereby the loading of samples and casting of gels.

Advantageously, the method according to the present invention may be combined with the method taught in our co-pending patent application of even date (Agents ref.: 75.57466) which uses PCR to introduce loop structures which provide a permanently attached 3' primer at the 3' terminal of a DNA strand of interest. For example, in such a modified method, the extension primer is introduced as part of the 3'-terminal loop structure onto a target sequence of one strand of double stranded DNA which contains the target position, said target sequence having a region A at the 3'-terminus thereof and there being optionally a DNA region B which extends 3' from region A, whereby said double-stranded DNA is subjected to polymerase chain reaction (PCR) amplification using a first primer hybridising to the 3'-terminus of the sequence complementary to the target sequence, which first primer is immobilised or provided with means for attachment to a solid support, and a second primer having a 3'-terminal sequence which hybridises to at least a portion of A and/or B of the target sequence while having at its 5'-end a sequence substantially identical to A, said amplification producing double-stranded target DNA having at the 3'-end of the target sequence, in the following order, the

region A, a region capable of forming a loop and a sequence A' complementary to sequence A, whereafter the amplified double-stranded DNA is subjected in immobilised form to strand separation whereby the non-immobilised target strand is liberated and region A' is permitted or caused to hybridise to region A, thereby forming said loop. The 3' end of region A' hybridises immediately adjacent the target position. The dideoxy and/or extension reactions use the hybridised portion as a primer.

The invention also comprises kits which will normally include at least the following components:

(a) a test specific primer which hybridises to sample DNA so that the target position is directly adjacent to the 3' end of the primer;

(b) a polymerase;

(c) enzyme means for identifying pyrophosphate release;

(d) deoxynucleotides; and

(e) optionally dideoxynucleotides.

If the kit is for use with initial PCR amplification then it will also normally include at least the following components:

(i) a pair of primers for PCR, at least one primer having means permitting immobilisation of said primer;

(ii) a polymerase which is preferably heat stable, for example TaqI polymerase;

(iii) buffers for the PCR reaction; and

(iv) deoxynucleotides.

Where an enzyme label is used to evaluate PCR, the kit will advantageously contain a substrate for the enzyme and other components of a detection system.

The invention will now be described by way of a non-limiting example with reference to the drawings in which:

Fig.1 shows a protocol for identifying a base in a single target position using the method according to the invention;

Fig.2 shows graphically the results obtained using the protocol shown in Fig.1.

Fig.3 shows a protocol for sequencing a sample DNA according one method of the invention;

Fig.4 shows graphically the results obtained using the protocol shown in Fig.3; and

Fig.5a and 5b show the relationship of response vs. DNA present and base length respectively.

Materials and Methods

Synthesis and purification of nucleotides

The oligonucleotides RIT320, RIT331, RIT332 and RIT333 (Table 1) complementary to a region in the active site of the HIV reverse transcriptase gene (encoding base 625 to 1165 see ref. 14) were synthesized by phosphoramidite chemistry in an automated DNA synthesis apparatus (Gene Assembler Plus, KABI Pharmacia AB, Sweden). Purification was performed with a protein liquid chromatography (FPLC) pepRPC 5/5 column (KABI Pharmacia AB, Sweden. PCR primers for amplification of cloned material were obtained from the Primer Set A solution from the Template Preparation Kit for DNA Sequencing (Dynal AS, Norway).

Table 1

The structure of the target gene and the primers for DNA mini-sequencing, cloning, and primer extension experiments.

Primer/template	Sequence 5'→ 3'	Used for
RIT 332	-CATCTGTTGAAGTGGGGACTT-	mini-sequencing and primer extension ELIDA
RIT 320 extension	-ATAAGAAAAAAGACAGTACTA-	primer ELIDA
univ. primer	-CCCGAATTCGATGGAGTTCATA ACCCATCCAAAG-	cloning of HIV RT fragment
RIT 333	-CCCGGATCCATACAATACTCCA GTATTTGC-	cloning of HIV RT fragment
HIV RT fragm. after the RIT 332 annealing site	-ATGTTTTTTGTCTGGTGTGGT-	mini-sequencing and primer extension ELIDA

Template preparation

A HIV reverse transcriptase gene fragment from a patient showing AZT resistance was PCR-cloned (Petterson, B, et al unpublished data) into the vector pRIT 28 (11) by using the primers RIT 331 and RIT 333. E.coli RR1ΔM15 was transformed and blue/white selectivity was used (15). PCR amplification was carried out by lysing a bacterial colony in 10 µl 20 mM Tris-Cl (pH 8.7) at 99 C for 5 minutes. Then 1 µl of the lysate was added to 5 pmol Primer set A, 200 µM dNTP, 20 mM Tris-Cl, pH 8.7, 2 mM $MgCl_2$, 0.1% Tween 20 and 0.5 units AmpliTaq DNA polymerase (Cetus, Ca., USA) making up a total volume of 50 µl. The temperature profile included a 0.5 min. denaturation step at 95°C

and a 1.5 min. annealing/extension step at 70°C, these steps were repeated 30 times. A GeneAmp PCR System 9600 (Perkin Elmer, Ca, USA) was used for both lysing the bacterial colony and running the reactions. The PCR product was immobilized on paramagnetic beads (16) with covalently coupled streptavidin, Dynabeads M280. The beads were used as described by the manufacturer (Dynal AS, Norway). Single stranded DNA was obtained by removing the supernatant after incubation of the immobilized PCR product in 0.10 M NaOH for 10 minutes. The immobilized single stranded DNA was washed with 50 µl 10 mM Tris-Cl (pH 7.5), 1 mM EDTA, 2 M NaCl, followed by 50 µl 10 mM Tris-Cl (ph 7.5). After washing, 20 mM Tris-Cl (pH 7.5), 8 mM $MgCl_2$ and 1 pmol sequencing primer were added to a final volume of 13 µl. The mixture was incubated at 65°C for 5 minutes and then cooled to room temperature.

Mini-sequencing

The dideoxynucleotide incorporation reactions were performed in a mixture of 1 µl (1/13 of a 50 µl PCR amplification reaction) of the template/primer-fragment immobilized on paramagnetic beads, 0.13 units Sequenase version 2.0 (United States Biochemical, USA), 0.5 µl 10 µM of a single ddNTP, and a buffer containing 25 mM Tris-Cl (pH 7.5), 12.5 mM $MgCl_2$ and 2.5 mM DTT in a final volume of 10 µl. After incubation at room temperature for 5 minutes, the beads were washed with 50 µl 10 mM Tris-Cl (pH 7.5), 1 mM EDTA, 2 M NaCl, 1% Tween 20 followed by 50 µl 10 mM Tris-Cl (pH 7.5), 1 mM EDTA, 2 M NaCl and finally with 50 µl 10 mM Tris-Cl (pH 7.5). The volume was adjusted to 5 µl with 10 mM Tris-Cl (pH 7.5). Control fragments were incubated with DNA polymerase in the absence of ddNTPS and zero control fragments in the presence of all ddNTPS. The different samples were subsequently analyzed with the ELIDA.

ELIDA

Samples from the above described mini-sequencing preincubation were assayed for full primer extension by the ELIDA. The assay was performed using an LKB 1250 luminometer and a potentiometric recorder. The luminometer was calibrated to give a response of 10 mV for the internal light standard. The luminescence output was calibrated by the addition of a known amount of ATP or ppi. The reaction was carried out at room temperature. The standard assay volume was 0.2 ml and contained the following components: 0.1 M Tris-acetate (pH 7.75), 2 mM EDTA, 10 mM magnesium acetate, 0.1% BSA, 1 mM DTT, 0.4 mg/ml polyvinylpyrrolidone 360,000, 2 μM dNTP, 100 μg/ml D-luciferin (BioOrbit, Finland), 4 μg/ml L-luciferin (BioOrbit, Finland), 0.3 units/ml ATP-sulfurylase (Sigma, USA) and purified luciferase (Enzymatix, UK). The amount of luciferase used gave a response of 1 V for 100 pmol ATP in a volume of 1 ml. After five minutes of preincubation, adenosine 5'-phosphosulfate, NaF and dNMP were added to final concentrations of 2 μM, 5 mM and 0.4 mM, respectively. The reaction was started after the addition of 5 μl of template/primer-fragments, taken from the dideoxy incorporation, by the addition of 0.13 units of Sequenase. The reaction was completed within 5 minutes.

Limited primer extension

After one mini-sequencing round it is possible to make a limited primer extension of the initial template/primer-fragment in the presence of the deduced complementary dNTP. 8 μl of the template/primer-fragment (8/13 of a 50 μl PCR amplification reaction) was preincubated for five minutes with 10 pmol of the deduced complementary dNTP, 0.26 units Sequenase, and 10 μl of a buffer containing 25 mM Tris-Cl (pH 7.5),

12.5 mM $MgCl_2$ and 2.5 mM DTT in a final volume of 20 µl. The beads were then washed as described above. The volume was finally adjusted to 15 µl with 10 mM Tris-Cl (pH 7.5). The new fragments were then used for a second round of mini-sequencing.

RESULTS

Principle of the mini-sequencing method

The principle of the mini-sequencing method is outlined in Fig. 1 in which the presence or absence of a T residue is investigated. The specific DNA-fragment of interest is amplified by PCR with one of the primers biotinylated in the 5' end. The PCR-amplified DNA is immobilized on magnetic beads containing covalently coupled streptavidin and subsequently converted into single stranded form by washing with NaOH, and a primer is annealed to the single stranded DNA. The template/primer-fragments are then divided into four different aliquots which are separately treated with one of the four ddNTPs in the presence of the polymerase. After the reaction, the resulting fragments are washed and used as substrate in a primer extension reaction with all four dNTPs present (see Fig. 1). The progress of the DNA-directed polymerisation reactions are monitored with the ELIDA. Incorporation of a dideoxynucleotide in the first reaction will prevent the formation of pyrophosphate during the subsequent "chase" reaction. In contrast, no dideoxynucleotide incorporation gives extensive pyrophosphate release during the "chase" reaction and this will lead to generation of light through the ELIDA reactions. From the ELIDA results, the first base after the primer is easily deduced. It is also possible to include both a negative control, which is incubated with all ddNTPs, and a positive control, which is incubated with DNA

polymerase in the absence of dNTPs.

Mini-sequencing of a specific DNA-fragment

As a model, a short stretch of the HIV-1 pol gene, which is responsible for increased resistance to AZT (Johan Wahlberg and M. Uhlen submitted), was chosen. The structure of the taget gene and the primers used in the experiments are shown in Table 1 above. Fig. 2 shows typical traces from ELIDA of the HIV-1 pol fragment. Incorporation of a single ddNTP was observed only when the complementary dideoxynucleotide (ddATP) was present during the polymerase reaction. No incorporation of noncomplementary bases was observed under the conditions used. The formation of ppi was detected by the ELIDA during the "chase" reaction only when template/primer-fragments were incubated with noncomplementary bases. When a complementary base was incorporated, no extension of the DNA was possible due to the lack of a free 3' OH group. The same result as above was obtained if the DNA-fragments (in the first step) were incubated with four different mixtures of three ddNTPs (not shown). It is important to note that a DNA polymerase lacking exonuclease activity must be used to obtain clean signals, although it is known that exonuclease activity of certain polymerases can be suppressed, e.g. by fluoride ions. It is also important to use low concentrations of nucleotides (0.05-5 μM) to avoid incorporation of non-complementary bases (data not shown).

Repeated mini-sequencing

After the first base was deduced, a limited incorporation of nucleotides with the complementary base (dATP) was performed. After washing, the template/primer-fragment was used for another sequencing step (Fig. 3, first row). The base following the dATP

fill-in reaction was deduced to be a G. The second row of Fig. 3 shows an additional mini-sequencing step of the template/primer-fragment obtained after a limited fill-in reaction in the presence of both dATP and dCTP. The next base in the sequence was determined to be C. In the last row, a limited polymerisation in the presence of dATP, dCTP and dGTP was performed. The following base was determined to be an A. Thus, four successive steps of mini-sequencing were performed with low background signal ineach step (see Figs. 3 and 4).

Sensitivity

In the experiments presented above 1/13th of a 50 µl PCR amplification reaction was used per ELIDA test. However, both lower and higher amounts can be used. In Fig. 5A, the initial rate and the extent of ppi formation during primer extension of a 161 bases long DNA-fragment as a function of DNA concentration is shown. Both the initial rate and the extent of ppi formed in the ELIDA are proportional to the DNA concentration in the interval tested (1/130 to 2/13 of a 50 µl PCR amplification reaction). The amount of DNA could be further increased as well as the binding capacity of the solid support to increase the signal of the assay. The upper limit for the assay (in a total volume of 200 µl) is 200 pmol ppi formed. The lower limit is mainly determined by the length of the DNA-fragment used, (as the signal is proportional to the amount of nucleotides incorporated during the primer extension reaction), the volume used and by contamination of ppi in the different solutions. Both these latter factors can be modified if necessary.

One important criteria for a successful on-line sequencer is the sensitivity of the assay and whether the method will be sensitive enough to measure the ppi formed after the incorporation of a single base. In Fig. 5B, the extent of ppi synthesis during primer

extension of DNA fragments of different lengths is shown as a function of the length of the DNA. The amount of ppi formed was proportional to the length of the fragment. Full primer extension of 1/13 of a 50 µl PCR amplification reaction of 161 bases long fragment gave a response of 110 mV.

REFERENCES

1. Sanger, F., Nicklen, S & Coulson, A.R. (1977) Proc Natl. Acad. Sci. USA 74, 5463-5467.
2. Maxam, A.M. & Gilbert, W. (1977) Proc. Natl. Acad. Sci. USA 74, 560-564.
3. Smith, L.M. Sanders, J.Z., Kaiser, R.J., Hughes, P. Dodd, C., Connell, C.R. Heiner, C.R., Kent, S.B.H. & Hood, L.E. (1986) Nature 321, 674-679.
4. Ansorge, W., Sproat, B., Stegemann, J., Schwager, C. & Zenke, M. (1987) Nucl. Acids Res. 15 4593-4602.
5. Driscoll, R.J., Youngquist, M.G. & Baldeschwieler, J.D. (1990) Nature 346 294-296.
6. Bains, W. & Smith, G.C. (1988) J. Theor. Biol. 135, 303-307.
7. Drmanac, R., Labat, I., Brukner, I. & Crkvenjakov, R. (1989) Genomics 4, 114-128.
8. Jett, J.H., Keller, R.A., Martin, J.C., Marrone, B.L. Moyzis, R.K., Ratliff, R.L., Seitzinger, N.K., Shera, E.B. & Stewart, C.C. (1989) Biomol. Struct. Dynamics 7, 301-306.
9. Antonarakis, S.E. (1989) N. Engl. J. Med. 320, 153-163.
10. Nigro, J.M., Baker, S.J. Preisinger, A.C., Jessup, J.M., Hostetter, R., Cleary, K., Bigner, S.H., Davidson, N., Baylin, S., Devilee, P., Glover, T., Collins, F.S., Weston, A., Modali, R., Harris, C.C. & Volgelstein, B. (1989) Nature 342, 705-708.
11. Hultman, T., Stahl, S., Moks, T. & Uhlén, M. (1988) Nucl. Acid Res. 17, 4937-4946.

12. Syvanen, A-C., Aalto-Setälä, K., Harju, Leena, Kontula, K. & Söderlund, H. (1990) Genomics 8, 684-692.

13. Nyrén, P. (1987) Anal. Biochem. 151, 235-238.

14. Myers, G., Korber. B., Berkofsky, J.A., Smith, R.F., and Pavlakis, G.N. Human Retroviruses and AIDS 1991 (Los Alamos National Laboratory, New Mexico 1991).

15. Rüther, U. (1982) Nucl. Acids Res. 10, 5765-5772.

16. Lea, T., Vartdal, F., Nustad, K., Funderud, S., Berge, A., Ellingsen, T., Schmid, R., Stenstad, P. & Ugelstad, J. (1988) J. Mol. Recognit. 1, 9-18.

17. Bessman, M.J., Lehman, I.R., Simms, E.S. & Kornberg, A. (1958) J Biol. Chem. 233, 171-177.

18. Adler, J., Lehman, I.R., Bessman, M.J., Simms, E.S. & Kornberg, A. (1958) Biochemistry 44, 641-647.

Claims

1. A method of identifying a base at a target position in a single-stranded DNA sequence wherein an extension primer, which hybridises to the sample DNA immediately adjacent to the target position is provided and the sample DNA and extension primer are subjected to a polymerase reaction in the presence of a deoxynucleotide or dideoxynucleotide whereby the deoxynucleotide or dideoxynucleotide will only become incorporated and release pyrophosphate if it is complementary to the base in the target position, any release of pyrophosphate being identified enzymically, different deoxynucleotides or dideoxynucleotides being added either to separate aliquots of sample-primer mixture or successively to the same sample-primer mixture and subjected to the polymerase reaction to indicate which deoxynucleotide or dideoxynucleotide is incorporated.

2. A method as claimed in claim 1 in which the said pyrophosphate is identified by reaction with luciferase and luciferin to generate a light signal.

3. A method as claimed in claim 1 or claim 2 in which the sample DNA is immobilised or provided with means for attachment to a solid support.

4. A method as claimed in claim 3 in which the sample DNA is prepared by amplification using at least one primer which is immobilised or provided with means for attachment to a solid support.

5. A method as claimed in any of claims 2-4 in which the light signal is determined quantitatively.

6. A method as claimed in claim 2 in which ATP is removed from the reaction mixture prior to the luciferase/luciferin reaction by contact with an

immobilised enzyme which converts ATP into a product which is not a substrate for luciferase.

7. A method as claimed in any of the preceding claims in which dideoxynucleotides are added in the absence of deoxynucleotides whereby the primer is extended only by the dideoxynucleotide complementary to the base in the target position.

8. A method as claimed in any of the preceding claims wherein the sample DNA is subjected to amplification; the amplified DNA is immobilised and then subjected to strand separation, the non-immobilised strand being removed and an extension primer is provided, which primer hybridises to the immobilised DNA immediately adjacent that portion of the DNA to be sequenced; each of four aliquots of the immobilised single stranded DNA is then subjected to a polymerase reaction in the presence of a deoxynucleotide, each aliquot using a different deoxynucleotide whereby only the deoxynucleotide complementary to the base in the target position becomes incorporated; pyrophosphate released by base incorporation being identified; the immobilised sample and primer then being separated from the reaction solution and the incorporated base added to the unreacted aliquots of sample/primer under polymerising conditions to extend the primer in all the aliquots by the said incorporated base and the immobilised sample/primer then being separated from the reaction solution, the process being repeated to sequence the sample DNA.

9. A method as claimed in any of the preceding claims in which the sample DNA is subjected to amplification; the amplified DNA is immobilised and then subjected to strand separation, the non-immobilised strand being removed and an extension primer is provided, which

primer hybridises to the immobilised DNA immediately adjacent that portion of the DNA to be sequenced; immobilised single stranded DNA is then subjected to a polymerase reaction in the presence of a first deoxynucleotide, and the extent of pyrophosphate release is determined, where necessary the immobilised sample and primer being separated from the reaction mixture and the reaction being repeated by successive addition of a second, third and fourth deoxynucleotide until a positive release of pyrophosphate indicates incorporation of a particular deoxynucleotide into the primer, whereupon the procedure is repeated to extend the primer one base at a time and to determine the base which is immediately 3'- of the extended primer at each stage.

10. A modification of the method as claimed in claim 8 or claim 9 wherein the said deoxynucleotide is a 3'-protected 2'-deoxynucleotide and after pyrophosphate release has been identified, the 3'-protecting group is removed prior to adding a further 3'-protected 2'-deoxynucleotide whereby repetition of this procedure is effected to sequence the sample DNA.

11. A kit for determining the base at a target position in sample DNA including at least the following components:

(a) a test specific primer which hybridises to sample DNA so that the target position is directly adjacent to the 3' end of the primer;

(b) a polymerase;

(c) enzyme means for identifying pyrophosphate release;

(d) deoxynucleotides; and

(e) optionally dideoxynucleotides.

12. A kit as claimed in claim 11 for use with initial PCR amplification including at least the following additional components:

(i) a pair of primers for PCR, at least one primer having means permitting immobilisation of said primer;

(ii) a polymerase which is preferably heat stable, for example Taq1 polymerase;

(iii) buffers for the PCR reaction; and

(iv) deoxynucleotides.

1/8

FIG. 1

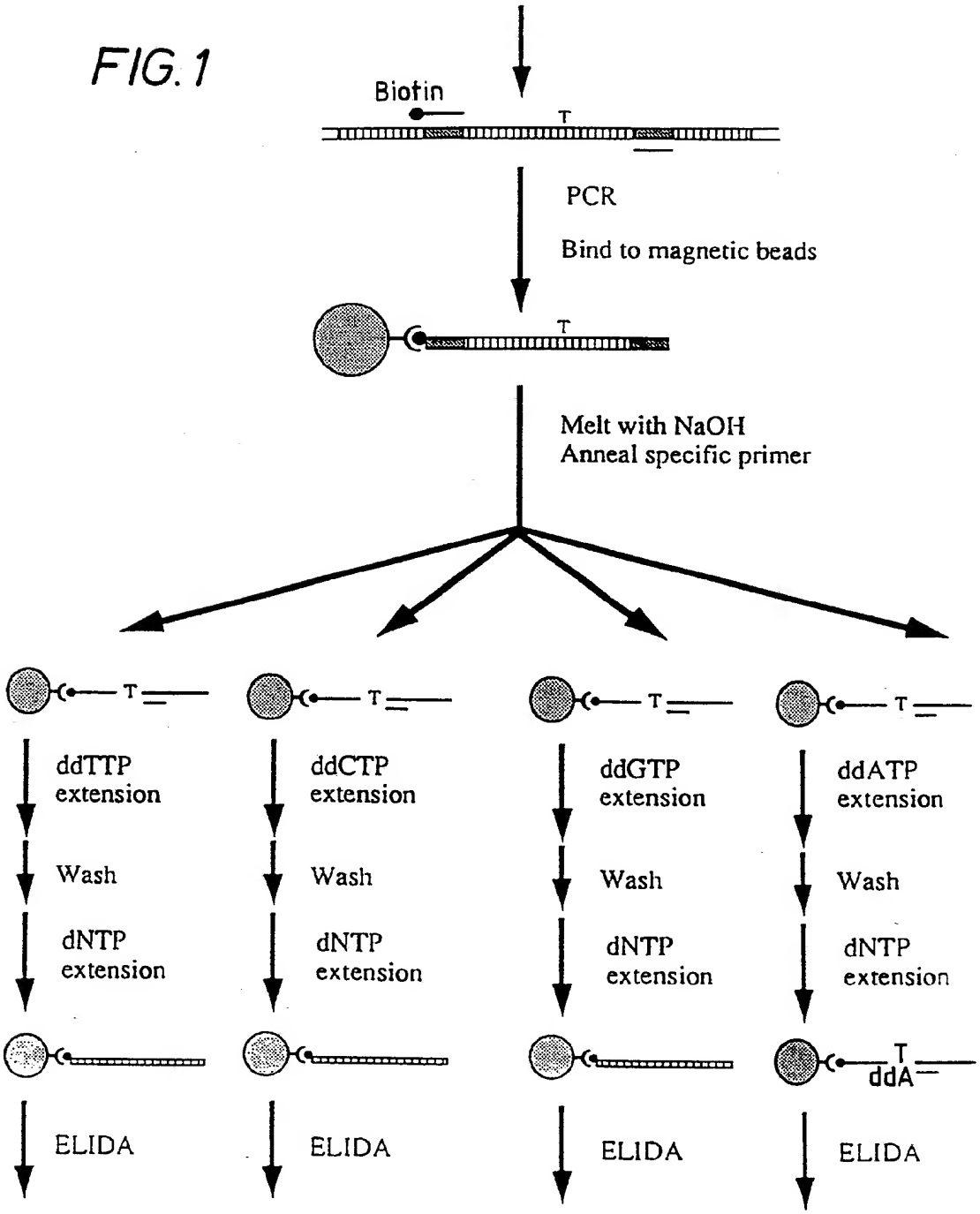

Genomic DNA / Cloned DNA
Biotin
T
PCR
Bind to magnetic beads
T
Melt with NaOH
Anneal specific primer
T
ddTTP extension
Wash
dNTP extension
ELIDA
T
ddCTP extension
Wash
dNTP extension
ELIDA
T
ddGTP extension
Wash
dNTP extension
ELIDA
T
ddATP extension
Wash
dNTP extension
T
ddA
ELIDA

FIG. 2

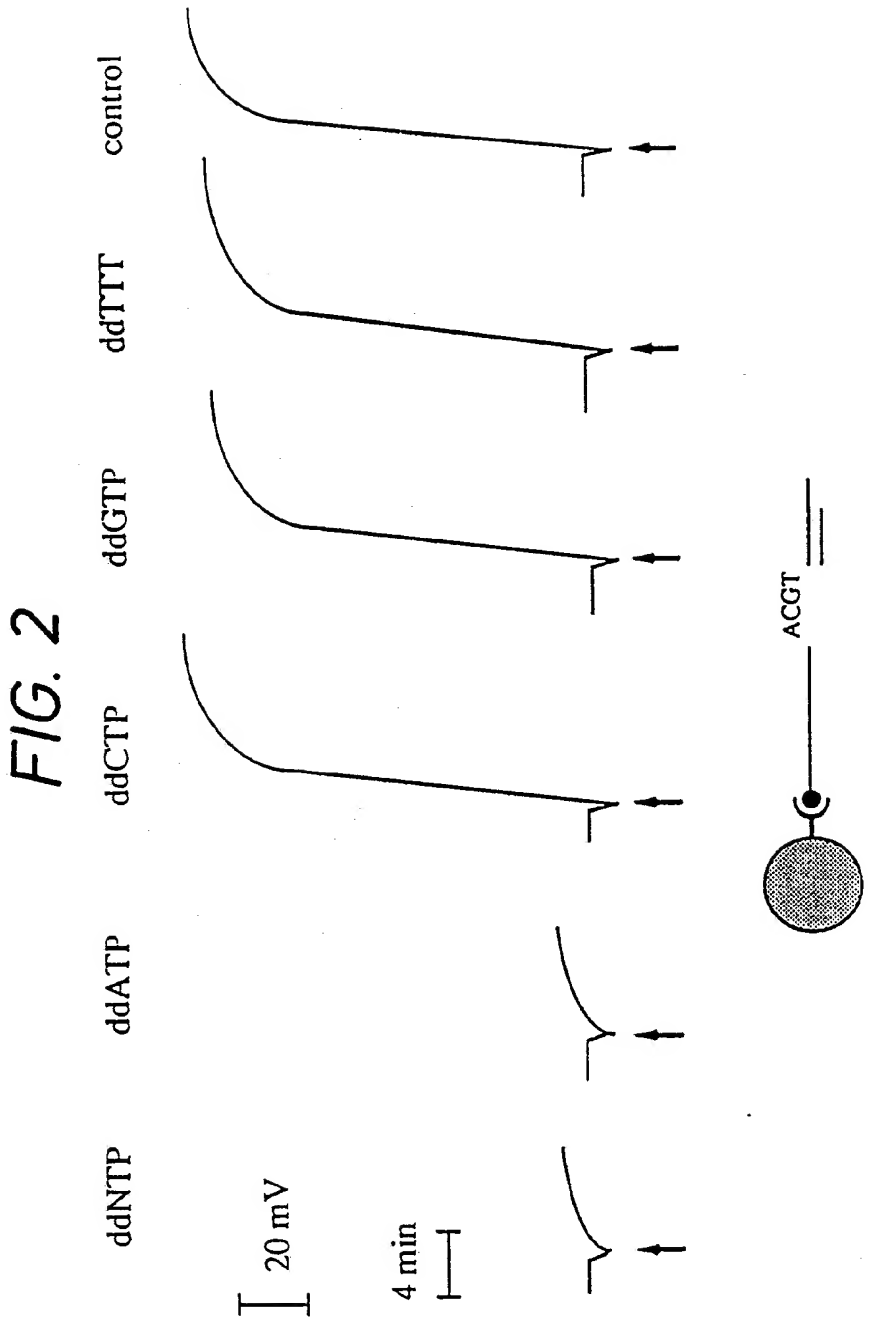


FIG. 3

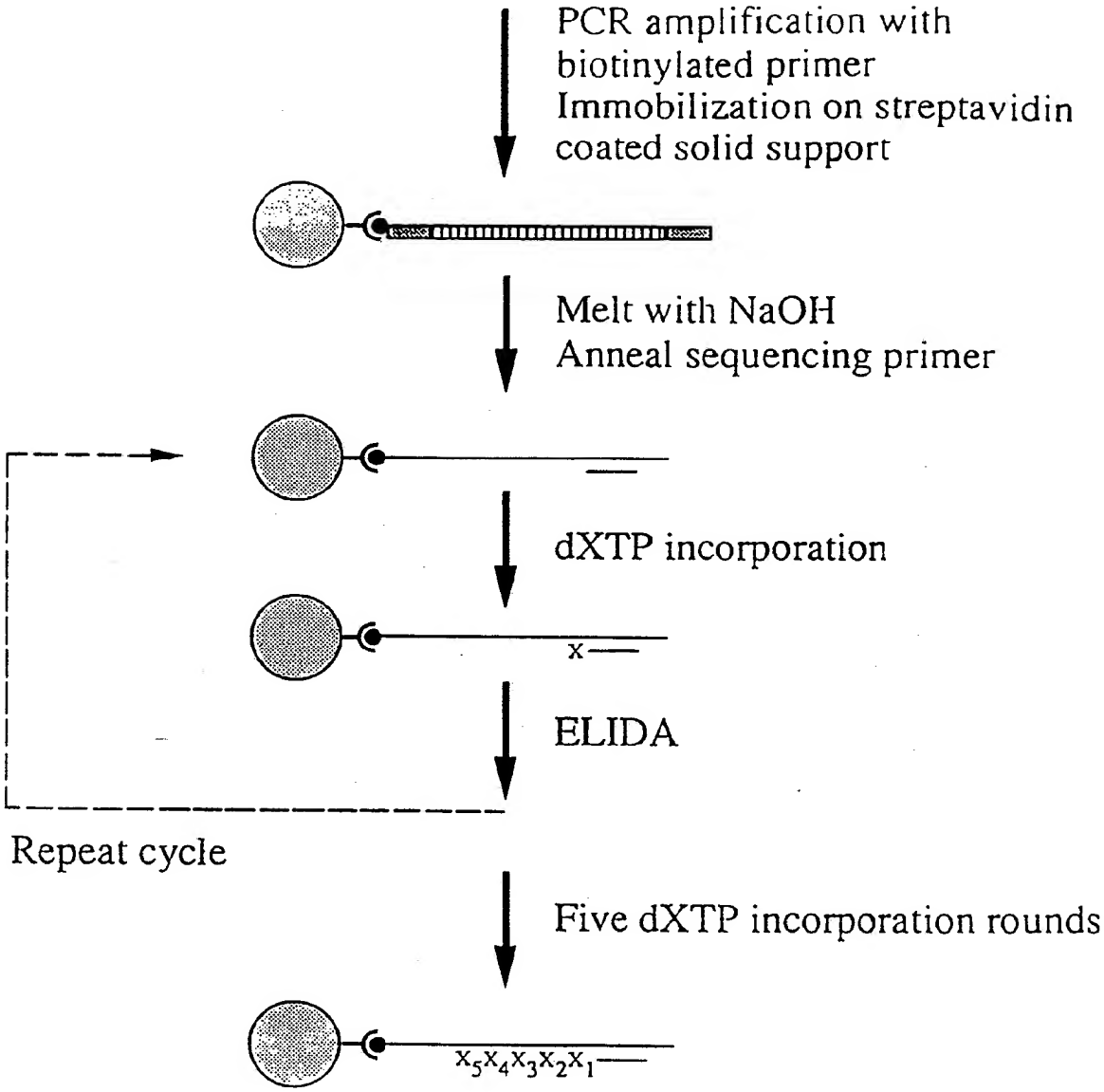

Genomic DNA / Cloned DNA
PCR amplification with
biotinylated primer
Immobilization on streptavidin
coated solid support
Melt with NaOH
Anneal sequencing primer
dXTP incorporation
x
ELIDA
Repeat cycle
Five dXTP incorporation rounds
$x_5x_4x_3x_2x_1$

FIG. 4 (1/3)

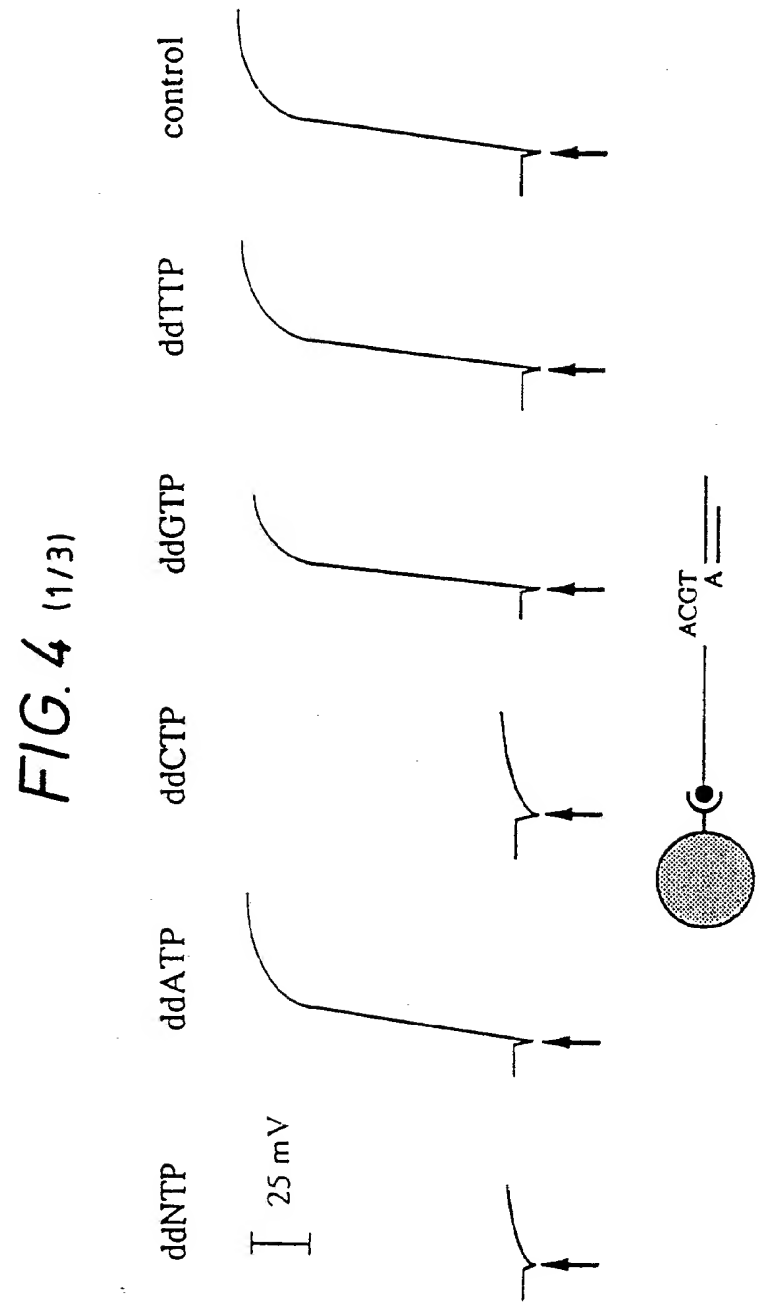
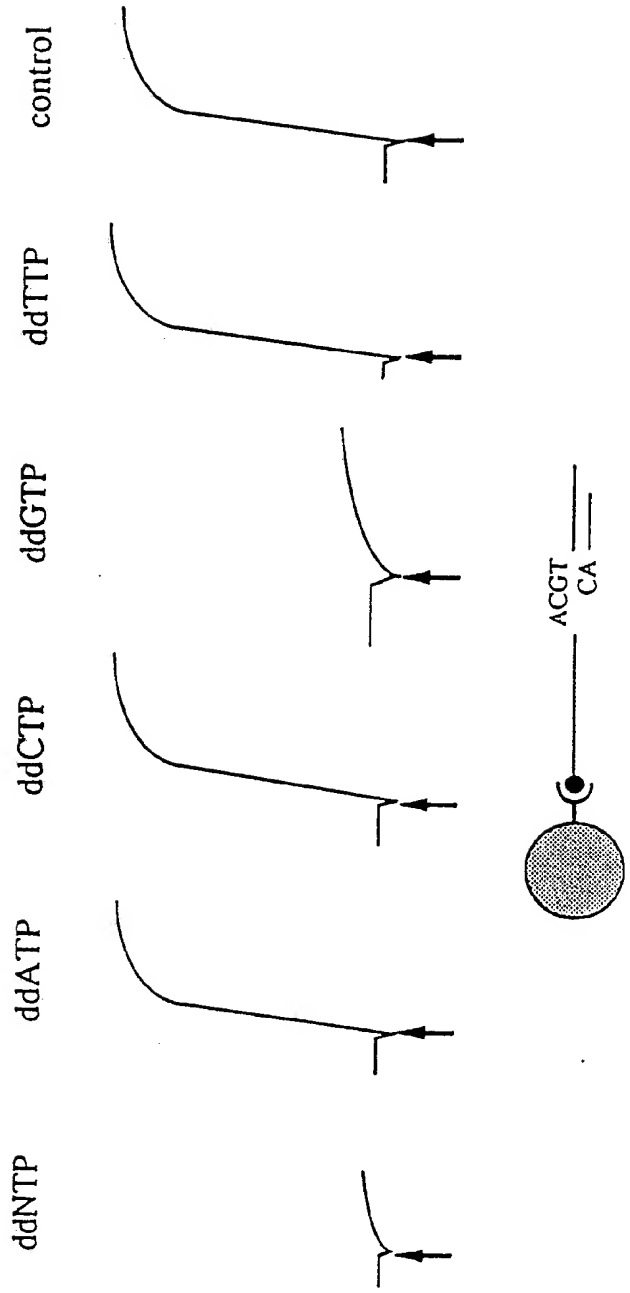


FIG. 4 (2/3)



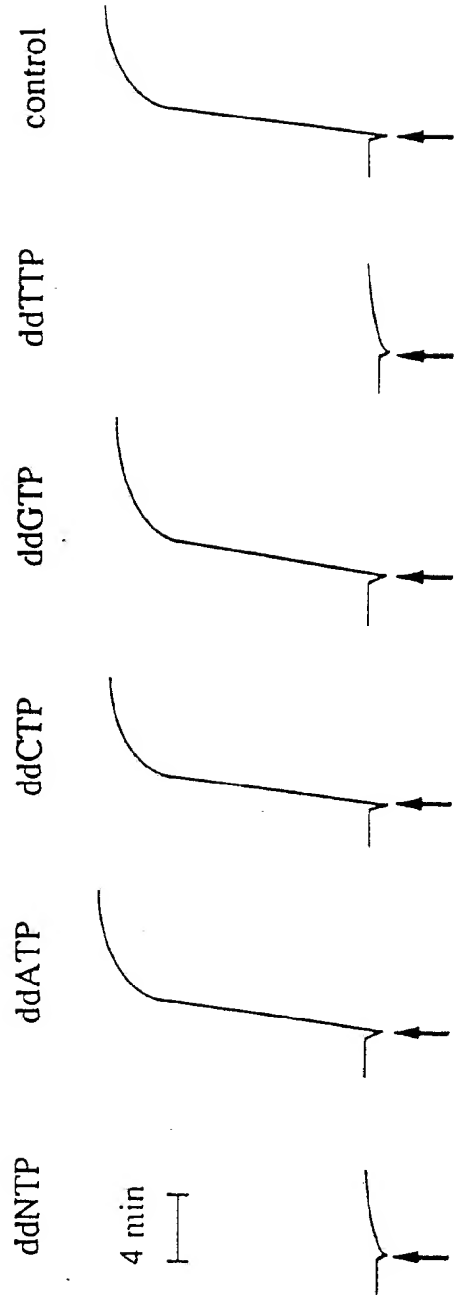


FIG. 4 (3/3)

FIG. 5a

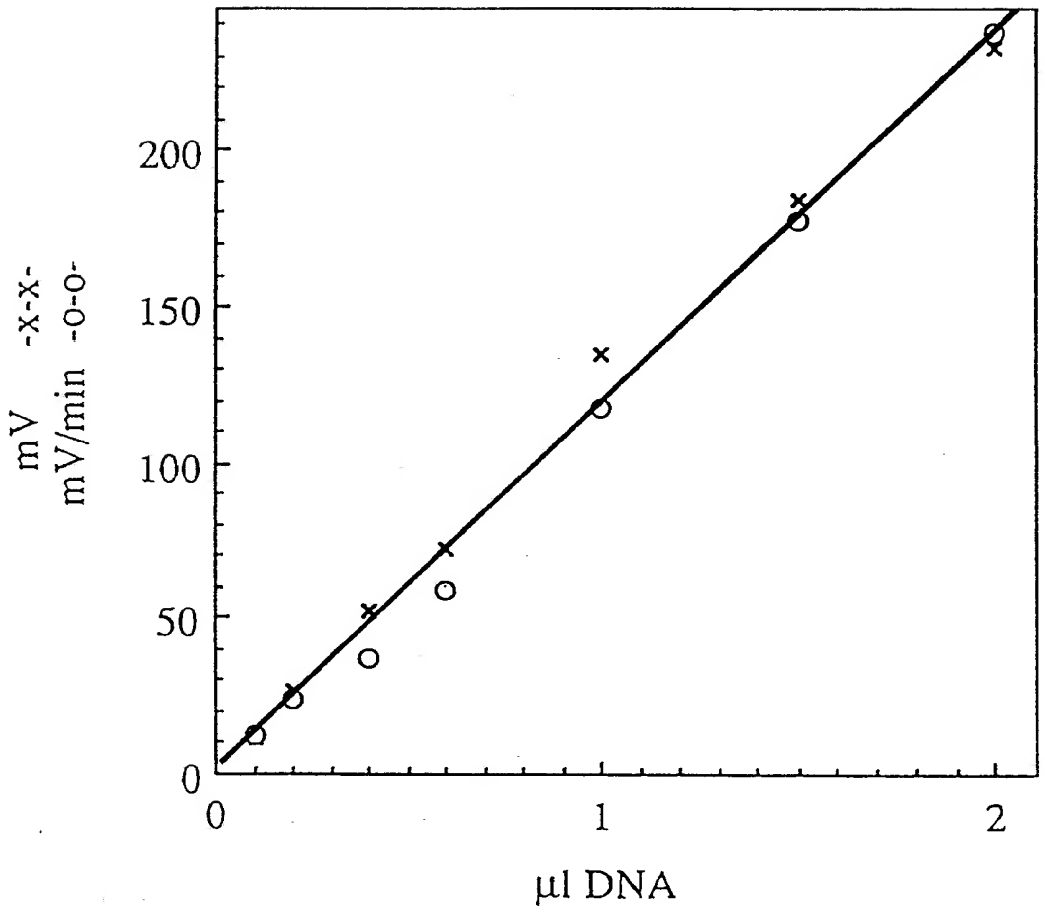
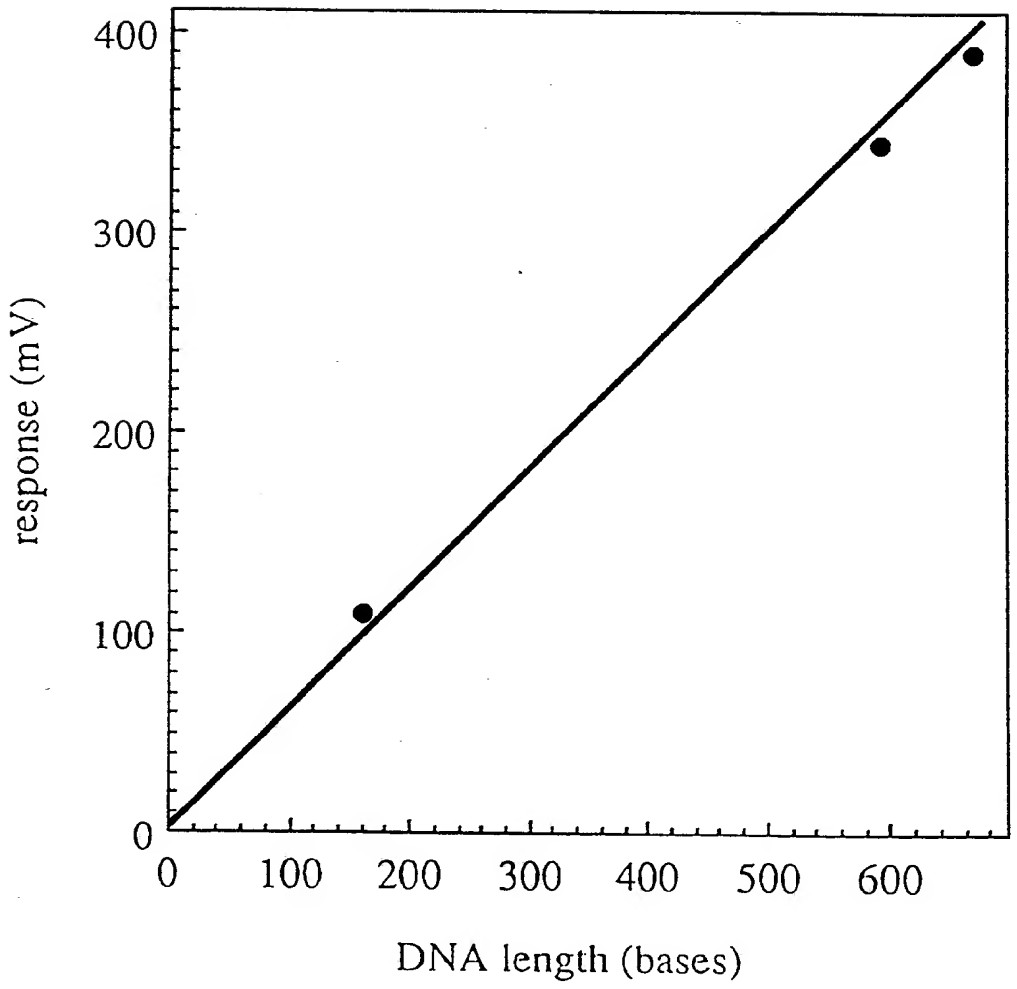


FIG. 5b


400
300
200
100
0
response (mV)
0
100
200
300
400
500
600
DNA length (bases)